

GenCore version 4.5  
 Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2001, 03:28:25 ; Search time 1190.53 Seconds

(without alignments)  
 6041.430 Million cell updates/sec

Title: US-09-521-640-2  
 Perfect score: 465  
 Sequence: 1 gggagatatcgtagttcac.....cccttcttgcataaacacn 465

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0  
 Seq.-ned: 1344157 seqs, 7733874588 residues

Word size : 30

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_ba1:\*

2: gb\_ba2:\*

3: gb\_ba3:\*

4: gb\_in1:\*

5: gb\_in2:\*

6: gb\_in3:\*

7: gb\_on:\*

8: gb\_ov:\*

9: gb\_dat1:\*

10: gb\_dat2:\*

11: gb\_ph:\*

12: gb\_p11:\*

13: gb\_p12:\*

14: gb\_p13:\*

15: gb\_p14:\*

16: em\_ba1:\*

17: em\_ba2:\*

18: em\_fun:\*

19: em\_htgo\_hum:\*

20: em\_htgo\_inv:\*

21: em\_htgo\_rod:\*

22: em\_htg\_hum1:\*

23: em\_htg\_hum2:\*

24: em\_htg\_hum3:\*

25: em\_htg\_hum4:\*

26: em\_htg\_hum5:\*

27: em\_htg\_hum6:\*

28: em\_htg\_hum7:\*

29: em\_htg\_num8:\*

30: em\_htg\_inv1:\*

31: em\_htg\_inv2:\*

32: em\_htg\_other:\*

33: em\_htg\_rod:\*

34: em\_hum1:\*

35: em\_hum2:\*

36: em\_hum3:\*

37: em\_hum4:\*

38: em\_hum5:\*

39: em\_hum6:\*

40: em\_in:\*

41: em\_om:\*

42: em\_or:\*

44: em\_ov:\*

45: em\_pat:\*

46: em\_ph:\*

47: em\_p1:\*

48: em\_ts:\*

49: em\_ro:\*

50: em\_sy:\*

51: em\_un:\*

52: em\_yi:\*

53: gb\_sts1:\*

54: qb\_sts2:\*

55: qb\_sts3:\*

56: qb\_sy:\*

57: qb\_un:\*

58: qb\_yi:\*

59: qb\_v12:\*

60: qb\_htg1:\*

61: qb\_htg2:\*

62: qb\_htg3:\*

63: qb\_htg4:\*

64: qb\_htg5:\*

65: qb\_htg6:\*

66: qb\_htg7:\*

67: qb\_htg8:\*

68: qb\_htg9:\*

69: qb\_htg10:\*

70: qb\_htg11:\*

71: qb\_htg12:\*

72: qb\_htg13:\*

73: qb\_htg14:\*

74: qb\_htg15:\*

75: qb\_htg16:\*

76: qb\_htg17:\*

77: qb\_htg18:\*

78: qb\_htg19:\*

79: qb\_htg20:\*

80: qb\_htg21:\*

81: qb\_htg22:\*

82: qb\_htg23:\*

83: qb\_htg24:\*

84: qb\_htg25:\*

85: qb\_pr1:\*

86: qb\_pr2:\*

87: qb\_pr3:\*

88: qb\_pr4:\*

89: qb\_pr5:\*

90: qb\_pr6:\*

91: qb\_pr7:\*

92: qb\_pr8:\*

93: qb\_pr9:\*

94: qb\_rol:\*

95: qb\_ro2:\*

96: qb\_in4:\*

97: qb\_pr10:\*

98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query No.	Score	Match	Length	DB	ID	Description
8	8	8	8	8	8	8	8
No matches found							
Search completed: November 7, 2001, 04:50:11							
Job time: 4906 sec.							



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

scoring table: OLIGO-NUC  
 Gappop 60.0 , Gapext 60.0  
~~seu--ned:~~ 730101 seqs, 313950809 residues  
 word size : 30

total number of hits satisfying chosen parameters:

minimum DB seq length: 0

maxium UB seq length: 2000000

post-processing: Listing first 45 summaries

database : N Geneset 0601 : \*

```
1: /sids1/gcgdata/geneseq/gcgc
```

```
3. /SIDS1_1/gcdata/geneseq/geneseq/NA1982.DAT:*
4. /SIDS1_1/gcdata/geneseq/geneseq/NA1983.DAT:*
5. /SIDS1_1/gcdata/geneseq/geneseq/NA1984.DAT:*
6. /SIDS1_1/gcdata/geneseq/geneseq/NA1985.DAT:*
7. /SIDS1_1/gcdata/geneseq/geneseq/NA1986.DAT:*
8. /SIDS1_1/gcdata/geneseq/geneseq/NA1987.DAT:*
9. /SIDS1_1/gcdata/geneseq/geneseq/NA1988.DAT:*
10. /SIDS1_1/gcdata/geneseq/geneseq/NA1989.DAT:*
11. /SIDS1_1/gcdata/geneseq/geneseq/NA1990.DAT:*
12. /SIDS1_1/gcdata/geneseq/geneseq/NA1991.DAT:*
13. /SIDS1_1/gcdata/geneseq/geneseq/NA1992.DAT:*
14. /SIDS1_1/gcdata/geneseq/geneseq/NA1993.DAT:*
15. /SIDS1_1/gcdata/geneseq/geneseq/NA1994.DAT:*
16. /SIDS1_1/gcdata/geneseq/geneseq/NA1995.DAT:*
17. /SIDS1_1/gcdata/geneseq/geneseq/NA1996.DAT:*
18. /SIDS1_1/gcdata/geneseq/geneseq/NA1997.DAT:*
19. /SIDS1_1/gcdata/geneseq/geneseq/NA1998.DAT:*
20. /SIDS1_1/gcdata/geneseq/geneseq/NA1999.DAT:*
21. /SIDS1_1/gcdata/geneseq/geneseq/NA2000.DAT:*
22. /SIDS1_1/gcdata/geneseq/geneseq/NA2001.DAT:*
```

Prep. No. is the number of results predicted by chance to have a scores greater than or equal to the score of the total score distribution.

## SUMMARIES

result	Score	Query No.	Match	Length	DB	ID	Description
--------	-------	-----------	-------	--------	----	----	-------------

卷之三

Search completed: November 7, 2001, 04:52:32  
Search time: 2005-02-05 10:45:20

Search completed: November 7, 2001, 04:52:32  
Search time: 2005-02-02



```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est710:*
164: gb_est721:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est75:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthnum29:*
177: em_esthnum30:*
178: em_esthnum31:*
179: em_esthnum32:*
180: em_esthnum33:*
181: em_estom3:*
182: em_estp11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*
190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_p1n1:*
208: em_gss_p1n2:*
209: em_gss_p1n3:*
210: em_gss_p1n4:*
211: em_gss_p1n5:*
212: em_gss_p1n6:*
213: em_gss_p1n7:*
214: em_gss_p1n8:*
215: em_gss_p1n9:*
216: em_gss_p1n10:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
No matches found					

Search completed: November 7, 2001, 04:30:08  
Job time: 3768 sec

Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2001, 03:24:30 ; Search time 1194.37 Seconds  
 (without alignments)  
 6022.006 Million cell updates/sec

Title: US-09-521-640-2  
 Perfect score: 465  
 Sequence: 1 ggaggatcgatcgatcac.....cccttcgttcaaaacach 465

Scoring table: IDENTIFY\_NUC  
 Gapop 10.0 , Gapext 1.0

Seq.:shed: 1344157 seqs, 733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : GenEmbl:  
 1: gb\_ba1:  
 2: gb\_ba2:  
 3: gb\_ba3:  
 4: gb\_in1:  
 5: gb\_in2:  
 6: gb\_in3:  
 7: gb\_on:  
 8: gb\_ov:  
 9: gb\_pat1:  
 10: gb\_pat2:  
 11: gb\_ph:  
 12: gb\_p11:  
 13: gb\_p12:  
 14: gb\_p13:  
 15: gb\_p14:  
 16: em\_ba1:  
 17: em\_ba2:  
 18: em\_fun:  
 19: em\_htgo\_hum:  
 20: em\_htgo\_inv:  
 21: em\_htgo\_rod:  
 22: em\_htg\_hum1:  
 23: em\_htg\_hum2:  
 24: em\_htg\_hum4:  
 25: em\_htg\_hum5:  
 26: em\_htg\_hum6:  
 27: em\_htg\_hum7:  
 28: em\_htg\_hum8:  
 30: em\_htg\_inv1:  
 31: em\_htg\_inv2:  
 32: em\_htg\_other:  
 33: em\_htg\_rod:  
 34: em\_hum1:  
 35: em\_hum2:  
 36: em\_hum3:  
 37: em\_hum4:  
 38: em\_hum5:  
 39: em\_hum6:  
 40: em\_hum7:  
 41: em\_ln:  
 42: em\_on:  
 43: em\_or:  
 44: em\_ov:  
 45: em\_pat:  
 46: em\_ph:  
 47: em\_p:  
 48: em\_ro:  
 49: em\_sts:  
 50: em\_un:  
 51: em\_v:  
 52: gb\_sts1:  
 53: gb\_sts2:  
 54: gb\_sts3:  
 55: gb\_sy:  
 56: gb\_un:  
 57: gb\_vil:  
 58: gb\_vil1:  
 59: gb\_vil2:  
 60: gb\_htg1:  
 61: gb\_htg2:  
 62: gb\_htg3:  
 63: gb\_htg4:  
 64: gb\_htg5:  
 65: gb\_htg6:  
 66: gb\_htg7:  
 67: gb\_htg8:  
 68: gb\_htg9:  
 69: gb\_htg10:  
 70: gb\_htg11:  
 71: gb\_htg12:  
 72: gb\_htg13:  
 73: gb\_htg14:  
 74: gb\_htg15:  
 75: gb\_htg16:  
 76: gb\_htg17:  
 77: gb\_htg18:  
 78: gb\_htg19:  
 79: gb\_htg20:  
 80: gb\_htg21:  
 81: gb\_htg22:  
 82: gb\_htg23:  
 83: gb\_htg24:  
 84: gb\_htg25:  
 85: gb\_pr1:  
 86: gb\_pr2:  
 87: gb\_pr3:  
 88: gb\_pr4:  
 89: gb\_pr5:  
 90: gb\_pr6:  
 91: gb\_pr7:  
 92: gb\_pr8:  
 93: gb\_pr9:  
 94: gb\_rol1:  
 95: gb\_f02:  
 96: gb\_in4:  
 97: gb\_pr10:  
 em\_ba3:  
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Length	DB	ID	Description
1	54.8	11	11918	65	AC018552 Homo sapi
2	52.6	11	7218	10	I66494 Sequence 14
c	3	37.2	8	43786	AC024859 Caenorhab
c	4	37.2	8	48896	AC006896 Caenorhab
c	5	36.8	7	485	AF309412 OncoflynC
c	6	36.4	7	215283	AC011966 Homo sapi
c	7	36.2	7	40552	AC084469 Caenorhab
c	8	36	7	74296	AC013805 Homo sapi





/codon\_start=1  
 /evidence-not\_experimental  
 /product="Hypothetical protein Y71H2AM. 8"  
 /protein\_id="AAK29971\_1"  
 /db\_xref="GI:13559767"  
 /translation="MAPSEATAELEDQILGKIAATLFEETLSKQPPSNTIVELMKHEVIR  
 RILNDIVSIOAKLUGDTDEIANSERDADSKTAKLAWAAEKFDTIGMDNVGLALFLACQ  
 LQANGDPIKSHIAYARPHAMTQLPTEPFLYDQQLKQSPFLEAFYFTTISROFCYFL  
 WIQNEBKSKQSYISLPLPTEPFLYDQQLKQSPFLEAFYFTTISROFCYFL  
 LAIAKNNKYEAAQERKDAANAMEPPIFVNPFVANEFPLKPYWAGVUTTRVNVMVP  
 ENQEVGHPVIFPLPMLVDPMANHAEVPLVTEPFLYDQQLKQSPFLEAFYFTTISROFCYFL  
 KAGNEVTFYGRSKEGHLLHNGPVTYHGKFDVLKLKIGIPKTDKTDKLLAKKKLQKF  
 VKKYCAAGNLFHDYLNCKNMRNISLKRIFC"  
 complement(16532 . 23598)  
 /gene="Y71H2AM.10"  
 CDS  
 complement(join(16732 . 17679,18063 . 18187,  
 18239 . 18398,19350 . 19840,20512 . 21131,21447 . 21202,  
 21749 . 21883,22158 . 22553,22316 . 23251,  
 23491 . 23598))  
 /gene="Y71H2AM.10"  
 /note="coded for by C. elegans cDNA yk285g2.5; coded for by C. elegans cDNA yk169a12.5; coded for by C. elegans cDNA yk205g2.3; coded for by C. elegans cDNA yk169a12.3"  
 /codon\_start=1  
 /product="Hypothetical protein Y71H2AM.10"  
 /protein\_id="AAK29969\_1"  
 /db\_xref="GI:13559765"  
 /translation="MATYSDLSQSSRRLSSAIEGKCYLNHRQVKSSCVIRPVPSSWEPVD  
 AVYDESLSHVGAICVQHQLSSTTAAKQDQHLLIEVNSRSLAKDQWANATESTVTLJSTAS  
 QNRTPSLVSYSVSEKDCDNFHDLQHLLIEVNSRSLAKDQWANATESTVTLJSTAS  
 QAFQNSMOTFLHVKSWANETQMEIRWMSRSNTIVPSNNWTRVSSASRPF  
 FVTSUVCPTERMTGENDNVAIILKMSSTDAKDDVILHNVILGKSSSEPPDDH  
 KVAQKFSVIADEVAIVVPKVTCGRSQVTFHAQKKTTSRHKELMLNVAISIQVH  
 SSRLRIFTISGAKRAEDVITQSHCSCDANTLQVQYAH  
 YLRITTTYSFRVNPKPKLPIKINMSSTSLTICKNKGFWRDPLGGYKRSRAOROFA  
 CWNRFQSHVVKLASLWIEDKEDQDLVLSHSISLLEDVNTVNTLQSNRIVANVK  
 FYEGRAKVGIVNGLAKLIVRNARKSDMSLVSITGSLARPICDTSFRI  
 LPLFLSPAFKVELAHSKSLTQYQSCDSFAVSYSSDATERPLNDLSSYFENSAHSD  
 NERALAVSHASHKVKVIADENNPLPVYEVISLOSSAQCASTINGASPARALVNLNP  
 ICKINNSPFSATDLSISFTLAPSISLPLHFLFISFVRRSAAF  
 KGYCOLVPPFSRSLSSGSSNSRDEBNWVLSQPOPSSSTISGGYGNKSTAERQS  
 SNGDDPSRSISIYGSESEVIAISQGNNVYVQSSRHPRTYLDNSDHNLLARIPKE  
 DRWTGHSQHOFHNTWKQRGGGGMAREPIRESIA"  
 complement(22704 . 33736)  
 /gene="Y71H2AM.11"  
 /note="coded for by C. elegans cDNA yk646h6.5; coded for by C. elegans cDNA yk646h6.3"  
 /codon\_start=1  
 /product="Hypothetical protein Y71H2AM.11"  
 /protein\_id="AAK29983\_1"  
 /db\_xref="GI:13559770"  
 /translation="MDQIDIAFKSISIDEDEFIDDLRESVATOSVADPARGDCVRM  
 SEWARDQDKLTGIVETSLWLGQQTPLPSGQOLPPIPAPAVGYYGRDKSKTLLTIGHLDV  
 QPAEKDGNTGNNPPELTDKLGQFEDPSGQOLPPIPAPAVGYYGRDKSKTLLTIGHLDV  
 VLECMEESESSGEGDGLDNEINDKISDVTPECSIDNYWLGKRNPKPCLPZALNDLWVMFS  
 QLWVYDQQLPGLGTAELVAPLTKDDEDEYKEIDCVDKNEKGSGHGMSDNEKNNLM  
 NRWRYSLSLSGVGEGASQPKGKSLRIVPHMPEATDKLNVSYLDLS  
 WAKERSPNFTFKVTSQHGNPWPWADERDANSAGRAIERDLENFRKHSKLLNFPPLK  
 YGMPPDTEFEGGSIPVTLTIQDLTKSPVMLPIGASDDMAHSNEKINRDNFYKGKVK  
 LAAVLFELAAH  
 35242 . 40205  
 /gene="Y71H2AM.7"  
 CDS  
 join(35242 . 35313,35416 . 35535,37225 . 37334,38048 . 38494,  
 39905 . 40097,40170 . 40205)  
 /codon\_start=1  
 /evidence-not\_experimental  
 /product="Hypothetical protein Y71H2AM.7"  
 /protein\_id="AAK29972\_1"  
 gene  
 complement(join(40206 . 41089,41240.41716 . 41854 ,  
 41904 . 42012,42332 . 42444,42498 . 42546))  
 /gene="Y71H2AM.12"  
 /note="coded for by C. elegans cDNA yk301b4.5; coded for by C. elegans cDNA yk301b4.5; coded for by C. elegans cDNA yk301b4.3"  
 /codon\_start=1  
 /product="Hypothetical protein Y71H2AM.12"  
 /protein\_id="AAK29982\_1"  
 /db\_xref="GI:13559778"  
 /translation="MDTAKTKVUVVVGESGAGKTAALLTCFLDNTFETDPLTTGIDFHK  
 HIKIVQLNDQSIQLWIDTAGQEFRQLAPAYIRSARAVLLVLDSDENCVEHLJRWK  
 GIIDKNNKDSITSTIVGKNDLVDSEKRPHTATRETNDEYIETSARMRKNIKKLFS  
 SVACRPFPHETPSOTILWNEPEPSAETKRCQCWN"  
 complement(42822 . 48202)  
 /gene="Y71H2AM.13"  
 complement(join(48822 . 43008,43641 . 43770,44200 . 44526,  
 44910 . 45066,45113 . 45297,46544 . 46593,47007 . 47099 ,  
 47964 . 48023,48089 . 48202))  
 /gene="Y71H2AM.13"  
 /note="coded for by C. elegans cDNA yk75h9.5; coded for by C. elegans cDNA yk47e3.5; coded for by C. elegans cDNA yk513a6.5; coded for by C. elegans cDNA yk100g10.5; coded for by C. elegans cDNA yk65c4.5; coded for by C. elegans cDNA yk19d11.5; coded for by C. elegans cDNA yk165d11.3; coded for by C. elegans cDNA yk23h4.3; coded for by C. elegans cDNA yk647e3.3; coded for by C. elegans cDNA yk75h9.3; coded for by C. elegans cDNA yk378b12.3"  
 /codon\_start=1  
 /product="Hypothetical protein Y71H2AM.13"  
 /protein\_id="AAK29968\_1"  
 /db\_xref="GI:13559766"  
 /translation="MGGVPSKHTPHPHPPSQRQLTSYGPPIEGRLLHKEQVDAFQG  
 IPVAAPIGNLRFALPOHEKWTVEYREKSFQARGIODHVLSPKTSPOSEDNTLN  
 FPPWPTPKNETGPEVILTHGGVDSADGMDTICNGKNGLNGPNTLQVWVTTIQRFLS  
 GFWWTGDSIPDVALHMDVFAULKWKVNGLNGPNTLQVWVTTIQRFLS  
 PVSDLFKVKIPMGNACSWATHPRPNACRRAOEQFVDEDMNTLQVWVKEFLP  
 DKFASALNEAVDFTKDPPLIGKDNFLFIPRVMEURKEAMPKPRLLGCAKSEGKV  
 Query Match 8 0% ; Score 37.2 ; DB 6 ; Length 143786;  
 Best Local Similarity 53.1% ; Pred. No. 2 ; 8 ;  
 Matches 78 ; Conservative 0 ; Mismatches 69 ; Indels 0 ; Gaps 0 ;  
 Qy 311 ttggcccttaataaaaaacttctttgtatcttggaaataactttccctgtttttcagca 376  
 Db 7918B TTTCATCAAAAAAAACCCAAATTTCATTTTTAAACAA 79129  
 Qy 377 aatggaaacccctcgctcngtcccttggccggccctttgcacgctttttggactgtta 436  
 Db 7912B AATTCACAAAAAAATCATTTTTTGTGAAATTTGGAAACTATTTCGATTTTT 79069  
 Qy 437 tcaaaaaaccccttctttgttcaaaaa 463  
 Db 7906B GAAAAAAATCAATTATCTCTGAAAAAA 79042  
 RESULT 4  
 AC006896/\_C  
 LOCUS AC006896 298960 bp DNA  
 DEFINITION Caenorhabditis elegans clone Y71H2X, \*\*\* SEQUENCING IN PROGRESS  
 ACCESSION AC006896  
 \*\*\* 15 unordered pieces.



TITLE Homo sapiens chromosome 15, clone RP11-285I14  
 JOURNAL Unpublished  
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barra,N., Beckerly,R., Boguslavsky,L., Boukigalter,B., Brown,A., Castile,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeAngelis,K., Dewar,K., Domino,M., Donegan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Halligan,J., Gardyna,S., Grant,G., Hagos,B., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehocky,J., Lieu,C., Locke,K., Macdonald,P., Marcus,N., McEwan,P., McGuirk,A., McErlean,K., McLaughlin,J., McNeil,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talanais,J., Tesfaye,S., Turrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.  
 COMMENT Direct Submission  
 Submitted (17-Oct-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Dec 23, 2000 this sequence version replaced gi:7960358.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L2605  
 Center clone name: 285.I.14  
 ----- Summary Statistics  
 Sequencing vector: M13; M77815; 47% of reads  
 Sequencing vector: Plasmid; n/a; 53% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.360731  
 Consensus quality: 193454 bases at least Q40  
 Consensus quality: 200496 bases at least Q30  
 Consensus quality: 205965 bases at least Q20  
 Insert size: 185000; agarose-fp  
 Insert size: 213083; sum-of-contigs  
 Quality coverage: 11.3 in Q20 bases; agarose-fp  
 Quality coverage: 9.8 in Q20  
 \* NOTE: This is a working draft/ sequence. It currently  
 \* consists of 23 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 1012: contig of 1012 bp in length  
 \* 1013 1112: gap of 100 bp  
 \* 1113 2148: contig of 1036 bp in length  
 \* 2149 2248: gap of 100 bp  
 \* 2249 3454: contig of 1206 bp in length  
 \* 3455 3554: gap of 100 bp  
 \* 3555 4747: contig of 1193 bp in length  
 \* 4748 4847: gap of 100 bp  
 \* 4848 5856: contig of 1009 bp in length  
 \* 5857 5956: gap of 100 bp  
 \* 5957 7144: contig of 1188 bp in length  
 \* 7145 7244: gap of 100 bp  
 \* 7245 8271: contig of 1027 bp in length  
 \* 8272 8371: gap of 100 bp  
 \* 8372 9382: contig of 1011 bp in length  
 \* 9383 9482: gap of 100 bp  
 \* 9483 10621: contig of 1139 bp in length  
 \* 10622 10721: gap of 100 bp  
 \* 10722 11851: contig of 1130 bp in length  
 \* 11852 13012: contig of 1065 bp in length  
 \* 11952 11151: gap of 100 bp  
 \* 11953 14692: gap of 100 bp  
 \* 14693 15322: contig of 1240 bp in length  
 \* 15933 16032: gap of 100 bp  
 \* 15934 17854: contig of 1822 bp in length  
 \* 17855 17954: gap of 100 bp  
 \* 17955 19826: contig of 1872 bp in length  
 \* 19827 19926: gap of 100 bp  
 \* 19927 21064: contig of 1138 bp in length  
 \* 21065 21164: gap of 100 bp  
 \* 21165 23108: contig of 1944 bp in length  
 \* 23109 23208: gap of 100 bp  
 \* 23209 24164: contig of 1556 bp in length  
 \* 24765 24864: gap of 100 bp  
 \* 24865 26747: contig of 1883 bp in length  
 \* 26748 26847: gap of 100 bp  
 \* 26848 28616: contig of 1759 bp in length  
 \* 28607 28706: gap of 100 bp  
 \* 28707 30429: contig of 1723 bp in length  
 \* 30430 30529: gap of 100 bp  
 \* 30530 32798: contig of 2251 bp in length  
 \* 32791 32880: gap of 100 bp  
 \* 32881 215383: contig of 182403 bp in length.  
 FEATURES source  
 1. .215383  
 LOCATION/Qualifiers  
 1. /clone="RP11-285I14"  
 /clone\_lib="RPCI-11 Human Male BAC"  
 /map="15"  
 /note="assembly\_fragment"  
 1. .1012  
 misc\_feature  
 1113. .2148  
 /note="assembly\_fragment"  
 misc\_feature  
 2249. .3454  
 /note="assembly\_fragment"  
 misc\_feature  
 3555. .4747  
 /note="assembly\_fragment"  
 misc\_feature  
 4848. .5856  
 /note="assembly\_fragment"  
 5957. .7144  
 /note="assembly\_fragment"  
 misc\_feature  
 7245. .8271  
 /note="assembly\_fragment"  
 misc\_feature  
 8372. .9382  
 /note="assembly\_fragment"  
 misc\_feature  
 9483. .10621  
 /note="assembly\_fragment"  
 misc\_feature  
 10722. .11851  
 /note="assembly\_fragment"  
 misc\_feature  
 11952. .13017  
 /note="assembly\_fragment"  
 misc\_feature  
 13118. .14592  
 /note="assembly\_fragment"  
 misc\_feature  
 14693. .15932  
 /note="assembly\_fragment"  
 misc\_feature  
 16033. .17854  
 /note="assembly\_fragment"  
 misc\_feature  
 17955. .19826  
 /note="assembly\_fragment"  
 misc\_feature  
 19927. .21064  
 /note="assembly\_fragment"  
 misc\_feature  
 21165. .23108  
 /note="assembly\_fragment"  
 misc\_feature  
 23209. .24764  
 /note="assembly\_fragment"  
 misc\_feature  
 24865. .26747  
 /note="assembly\_fragment"  
 misc\_feature  
 26848. .28616  
 /note="assembly\_fragment"  
 misc\_feature  
 28607. .28706  
 /note="assembly\_fragment"  
 misc\_feature  
 28707. .30429  
 /note="assembly\_fragment"  
 misc\_feature  
 30430. .30529  
 /note="assembly\_fragment"  
 misc\_feature  
 30530. .32798  
 /note="assembly\_fragment"  
 misc\_feature  
 32791. .32880  
 /note="assembly\_fragment"  
 misc\_feature  
 32881. .215383  
 /note="assembly\_fragment"





TITLE	JOURNAL	REFERENCE
and Gibbs, R.	Direct Submission	
	Unpublished	
	2 (bases 1 to 33)	

AUTHORS Worley, K.C.

FILE DIRECT SUBMISSION SUBMITTED 6/22/2013

JOURNAL SUBMITTED (23-MAR-OF MOLECULAR AND

Dr. Moretti and  
Baylor Plaza: How

COMMENT On Sep 10, 2000 t

- Ge

Center: Baylo

Center code: 112233445566778899

WEB SITE: <http://www.elliptical.com>

*	244464	247852;	contig of 3389 bp in length	Qy	347	tttcggaaataaactttccgttttcagcaaataatggaaaccctgtcttggc 406	
*	247952	gap of unknown length	Db	109725	ATGTCACTATGTTTCACTATCTCGCCCTTGCAATGCTGTCATGGA 109784		
*	251204;	contig of 3232 bp in length					
*	251305	gap of unknown length					
*	253924;	contig of 2620 bp in length	Qy	407	99gccttgcacgtttggactgttatcaaaaaacccttc 451		
*	254024	gap of unknown length	Db	109785	CAGCTTTCTTCAGTTCGACTGGAAATCCCTAAATCCATC 109829		
*	258689	contig of 4665 bp in length					
*	258690	gap of unknown length					
*	258720	contig of 2566 bp in length					
*	261356	gap of unknown length					
*	261456	263987;	contig of 2532 bp in length	RESULT	10		
*	263988	264087;	gap of unknown length	AL365272	132449 bp	DNA	
*	264088	gap of unknown length	LOCUS	AL365272	Human DNA sequence from clone RP11-328C17 on chromosome 6, complete		
*	267442	267591;	contig of 3040 bp in length	DEFINITION			
*	267592	269625;	gap of unknown length	ACCESSION	AL365272		
*	269626	269725;	contig of 2034 bp in length	KEYWORDS	HTG,		
*	269726	271829;	gap of unknown length	SOURCE	Homo sapiens		
*	271830	271929;	gap of unknown length	ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. human.		
*	271930	274961;	contig of 3032 bp in length	COMMENT	Health, P.		
*	274962	275061;	gap of unknown length	AUTHORS	Direct Submission		
*	275062	277232;	contig of 2171 bp in length	TITLE	CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk		
*	277233	277332;	gap of unknown length	JOURNAL	requests: clonequery@sanger.ac.uk		
*	277333	278978;	contig of 2104 bp in length	COMMENT	On Dec 24, 2000 this sequence version replaced gi:11875910.		
*	278979	279078;	gap of unknown length	AUTHORS	During sequence assembly data is compared from overlapping clones.		
*	279079	281944;	contig of 2866 bp in length	TITLE	Where differences are found these are annotated as variations		
*	281945	282044;	gap of unknown length	JOURNAL	together with a note of the overlapping clone name. Note that the		
*	282045	284703;	contig of 2659 bp in length	COMMENT	variation annotation may not be found in the sequence submission		
*	284704	284803;	gap of unknown length	AUTHORS	corresponding to the overlapping clone, as we submit sequences with		
*	284804	286499;	contig of 1646 bp in length	TITLE	only a small overlap as described above.		
*	286500	286599;	gap of unknown length	JOURNAL	This sequence has been finished according to sequence map criteria		
*	286600	289303;	contig of 2704 bp in length	COMMENT	as follows. An attempt is made to resolve all sequencing problems,		
*	289304	289403;	gap of unknown length	AUTHORS	such as compressions and repeats, but not necessarily within known		
*	289404	291372;	contig of 1969 bp in length	TITLE	annotated repeat sequence elements. Where the sequence is		
*	291373	291472;	gap of unknown length	JOURNAL	ambiguous, there is an annotation using the 'unsure' feature key.		
*	291473	293291;	contig of 1821 bp in length	COMMENT	The following abbreviations are used to associate primary accession		
*	293294	293393;	gap of unknown length	AUTHORS	numbers given in the feature table with their source databases:		
*	293394	295064;	contig of 1671 bp in length	TITLE	Em:, EMBL:, Sw:, SWISSPROT;, Tr:, TREMBL;, Wp:, WORMPEP;, Information		
*	295065	295164;	gap of unknown length	JOURNAL	on the WORMPEP database can be found at		
*	295165	296870;	contig of 1706 bp in length	COMMENT	<a href="http://www.sanger.ac.uk/Projects/C_elegans/wormpep">http://www.sanger.ac.uk/Projects/C_elegans/wormpep</a> This sequence		
*	296871	298567;	contig of 1821 bp in length	AUTHORS	was generated from part of bacterial clone contigs of human		
*	298569	298668;	gap of unknown length	TITLE	chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping		
*	298669	299734;	contig of 1666 bp in length	JOURNAL	Group. Further information can be found at		
*	299735	299834;	gap of unknown length	COMMENT	<a href="http://www.sanger.ac.uk/Chori.org/bacpac/home.htm">http://www.sanger.ac.uk/Chori.org/bacpac/home.htm</a>		
*	299835	301632;	contig of 1798 bp in length	AUTHORS	IMPORTANT: This sequence is not the entire insert of clone		
*	301633	301734;	gap of unknown length	TITLE	RP11-328C17. It may be shorter because we sequence overlapping		
*	301733	303247;	contig of 1598 bp in length	JOURNAL	sections only once, except for clone RP11-328C17 is at 132449.		
*	303248	303341;	gap of unknown length	COMMENT	The true right end of clone RP11-328C17 is at 132449. In this		
*	303348	304660;	contig of 1313 bp in length	AUTHORS	sequence. The true left end of clone RP11-328C17 is at 64833 in this		
*	304661	304760;	gap of unknown length	TITLE	sequence. The true right end of clone RP1-125A24 is at 100 in this		
*	304761	306434;	contig of 1694 bp in length	JOURNAL	sequence. The true right end of clone RP1-125A24 is at 61243 in		
*	306455	306551;	gap of unknown length	COMMENT	this sequence.		
*	306555	307630;	contig of 1515 bp in length	AUTHORS	FEATURES	LocationQualifiers	
*	307631	307730;	gap of unknown length	TITLE	1..132449		
*	307731	309334;	contig of 1604 bp in length	JOURNAL	/organism="Homo sapiens"		
*	309335	309434;	gap of unknown length	COMMENT	/chromosome="6"		
*	309435	310606;	contig of 1172 bp in length	AUTHORS	/clone="RP11-328C17"		
*	310607	310706;	gap of unknown length	TITLE	/clone="RP11-328C17"		
*	311845	311844;	contig of 1138 bp in length	JOURNAL	1..1562		
*	311845	313446;	gap of unknown length	COMMENT	/more="LIM repeat: matches 4133..5719 of consensus"		
*	313447	315156;	contig of 1710 bp in length	AUTHORS	1563..1641		
*	315157	315256;	gap of unknown length	TITLE	/more="LIM repeat: matches 1..80 of consensus"		
*				JOB	1669..2199		
*				repeat_region	/note="LIM4 repeat: matches 3396..3937 of consensus"		
*				repeat_region	/note="LIM4 repeat: matches 3396..3937 of consensus"		

Query Match 7 7% Score 36; DB 70; Length 334796;  
 Best Local Similarity 50.9%; Pred. No. 6; Gaps 0;  
 Matches 84; Conservative 0; Mismatches 81; Indels 0;  
 Gaps 0;  
 Query 287 tatacccttacacccctggacaaaccttggccatataaactattttgcata 346  
 Db 109665 TATCTCATTCCTCACCGTTACAAACTTGTTCAGGGATAGGCTCA 109724

```
repeat_region          2200. .2474  
/note="AlusX repeat: matches 38. .312 of consensus"  
repeat_region          4279 .4279  
/note="L1M4 repeat: matches 1686. .3396 of consensus"  
repeat_region          4589 .4589  
/note="L1M4 repeat: matches 1254. .1544 of consensus"  
repeat_region          4799 .4799  
/note="L1M4 repeat: matches 785. .924 of consensus"  
repeat_region          complement(4809. .5326)  
/note="match: GSS: Em: AQ774852"  
repeat_region          complement(4860. .5272)  
/note="match: GSS: Em: AQ787689"  
repeat_region          4943. .5094  
/note="L1M4C repeat: matches 272. .420 of consensus"  
repeat_region          complement(5020. .5329)  
/note="match: GSS: Em: AQ494601"  
repeat_region          5142. .5435  
/note="L1M1D repeat: matches 170. .428 of consensus"  
repeat_region          5744 .5744  
/note="AlusX repeat: matches 1. .310 of consensus"  
repeat_region          5745. .5869  
/note="L1M1D repeat: matches 428. .568 of consensus"  
repeat_region          6450. .6686  
/note="M1R repeat: matches 7. .261 of consensus"  
repeat_region          6916. .7994  
/note="L1M4 repeat: matches 3582. .4668 of consensus"  
repeat_region          7904. .8050  
/note="L1MB8 repeat: matches 5808. .5884 of consensus"  
repeat_region          8066. .8169  
/note="L1M4 repeat: matches 3479. .3593 of consensus"  
repeat_region          8178. .8224  
/note="L1MB4 repeat: matches 6124. .6169 of consensus"  
repeat_region          8222. .8525  
/note="AlusX repeat: matches 1. .301 of consensus"  
repeat_region          8520. .8915  
/note="L1MB4 repeat: matches 5722. .6124 of consensus"  
repeat_region          complement(8862. 9353)  
/note="match: GSS: Em: AQ355497"  
repeat_region          9029. .9150  
/note="2 copies 61 mer 95% conserved"  
repeat_region          misc_feature  
/note="match: GSS: Em: AQ175845"  
repeat_region          misc_feature  
/note="match: GSS: Em: AQ135206"  
repeat_region          misc_feature  
/note="match: GSS: Em: AQ717684"  
repeat_region          misc_feature  
/note="match: GSS: Em: AQ147759"  
repeat_region          misc_feature  
/note="MER53 repeat: matches 1. .1189 of consensus"  
repeat_region          10046. .10143  
/note="2 copies 49 mer 91% conserved"  
repeat_region          11168. .11361  
/note="L1T1B repeat: matches 101. .309 of consensus"  
repeat_region          11418. .11756  
/note="AluY repeat: matches 1. .309 of consensus"  
repeat_region          12109. .12515  
/note="L12 repeat: matches 1747. .2155 of consensus"  
repeat_region          12948. .12947  
/note="N1R repeat: matches 29. .206 of consensus"  
repeat_region          12915. .13026  
/note="N1R repeat: matches 186. .237 of consensus"  
repeat_region          13005. .13103  
/note="L2 repeat: matches 2417. .2516 of consensus"  
repeat_region          13331. .13664  
/note="L1T1B repeat: matches 4. .361 of consensus"  
repeat_region          13704. .13907  
/note="L1T33 repeat: matches 297. .511 of consensus"  
repeat_region          15450. .15530  
/note="L2 repeat: matches 2668. .2749 of consensus"  
repeat_region          15876. .15935
```

```

repeat_region /note="#2 copies 30 mer 100% conserved"
  16202 .16294
  /note="MFR81 repeat: matches 2. .94 of consensus"
  17335 .17668
  /note="match: GSS: Em:AQ104665"
  17419 .17715
  /note="AluSx repeat: matches 1. .296 of consensus"
  17905 .18362
  /note="match: GSS: Em:AQ151578"
  19374 .19496
  /note="L2 repeat: matches 2579. .2709 of consensus"
  20271 .20430
  /note="MFR20 repeat: matches 55. .218 of consensus"
  21095 .21389
  /note="AluSx repeat: matches 1. .295 of consensus"
  21840 .21946
  /note="MIR repeat: matches 42. .152 of consensus"
  22278 .22772
  /note="L2 repeat: matches 2242. .2710 of consensus"
  22758 .22848
  /note="L2 repeat: matches 2661. .2750 of consensus"
  23314 .24204
  /note="C/G Island"
  /note="not experimental"
  23516 .23601
  /note="Weak data."
  25664 .25783
  /note="L2 repeat: matches 2588. .2741 of consensus"
  25736 .26131
  /note="MFR3 repeat: matches 1. .324 of consensus"
  26858 .27183
  /note="AluJb repeat: matches 1. .301 of consensus"
  27208 .27334
  /note="MIR repeat: matches 81. .221 of consensus"
  27673 .27833
  /note="MIR repeat: matches 7. .175 of consensus"
  complement(28112 .28516)
  /note="match: GSS: Em:AQ212338"
  28535 .29088
  /note="match: GSS: Em:BB51862"
  28560 .28684
  /note="MIR repeat: matches 79. .210 of consensus"
  28518 .28828
  /note="match: GSS: Em:B78941"
  28588 .28841
  /note="match: GSS: Em:BB88272"
  28849 .29048
  /note="L2 repeat: matches 2. .214 of consensus"
  30148 .30232
  /note="L2 repeat: matches 2610. .2692 of consensus"
  complement(31265 .31862)
  /note="match: GSS: Em:AO625926"
  complement(31333 .31862)
  /note="match: GSS: Em:AO625934"
  complement(31491 .31857)
  /note="match: GSS: Em:AO131221"
  31733 .31875
  /note="L2 repeat: matches 2039. .2197 of consensus"
  33899 .33992
  /note="L2 repeat: matches 2578. .2667 of consensus"
  34111 .34427
  /note="MER31A repeat: matches 4. .477 of consensus"
  35300 .35300
  /note="MER5A repeat: matches 121. .189 of consensus"
  35435 .35435

Query Match Score 35.8; DB 90; Length 132449
Best Local Similarity: 7.7%; Pred. No. 7.4;
Matches 79; Mismatches 0; Deletions 0;
Query 214 ggccctg!; Subject 115574 GGCCCTG!; Length 132449

```



Center, A.F.A. & Green, P. (1996-1997)  
<http://fpc.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L9002  
 Center clone name: 178\_E\_20  
 Sequencing vector: M13; M77815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 156689 bases at least Q40  
 Consensus quality: 164477 bases at least Q30  
 Consensus quality: 167863 bases at least Q20  
 Insert size: 171000; agarose-fp  
 Insert size: 171010; sum-of-contigs  
 Quality coverage: 3.7 in Q20 bases; agarose-fp  
 Quality coverage: 3.7 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	119: contig of 19 bp in length
20	119: gap of 100 bp
120	1921: contig of 1802 bp in length
1922	2021: gap of 100 bp
2022	3689: contig of 1668 bp in length
3689	3789: gap of 100 bp
3790	6531: contig of 2742 bp in length
6532	6631: gap of 100 bp
6632	8134: contig of 1503 bp in length
8135	8234: gap of 100 bp
8235	11919: contig of 3685 bp in length
11920	12019: gap of 100 bp
12020	15244: contig of 3225 bp in length
15245	15344: gap of 100 bp
15345	117344: contig of 2000 bp in length
17345	17444: gap of 100 bp
17445	21830: contig of 4386 bp in length
21830	21930: gap of 100 bp
21931	25081: contig of 3151 bp in length
25082	25181: gap of 100 bp
25182	2529736: contig of 4555 bp in length
29737	29436: gap of 100 bp
29837	36853: contig of 7017 bp in length
36854	36953: gap of 100 bp
36954	414120: contig of 4317 bp in length
41271	41370: gap of 100 bp
41371	446765: contig of 5395 bp in length
46766	46665: gap of 100 bp
46866	57601: contig of 10736 bp in length
57602	57701: gap of 100 bp
57702	67879: contig of 10178 bp in length
67810	6779: gap of 100 bp
5	78235: contig of 10256 bp in length
7	78335: gap of 100 bp
7b	89176: contig of 10841 bp in length
8c	89276: gap of 100 bp
891	106051: contig of 16805 bp in length
891	10651: gap of 100 bp
10651	1181: gap of 100 bp
105	121387: contig of 15026 bp in length
123	124487: gap of 100 bp
123	137207: contig of 15720 bp in length
137	137207: gap of 100 bp
137	53745: contig of 16438 bp in length

FEATURES \* 153746 153845: 9ap of 100 bp  
Location/Qualifiers 153846 172307 Contig of 18462 bp in length.

source 1. .172307  
/organism="Homo sapiens"  
/db\_xref="taxon: 9606"  
/chromosome="2"  
/map="2"  
/clone="RP11-178E20"  
1. 19 /clone\_lib="RPCI-11 Human Male BAC"  
/note="assembly\_fragment  
clone\_end:T7  
vector\_side:left"  
120. .1921

misc\_feature /note="assembly\_fragment"  
2022. .36899  
misc\_feature /note="assembly\_fragment"  
3790. .65311  
misc\_feature /note="assembly\_fragment"  
6632. .81344  
misc\_feature /note="assembly\_fragment"  
8235. .11919  
misc\_feature /note="assembly\_fragment"  
12020. .152244  
misc\_feature /note="assembly\_fragment"  
15345. .17344  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:right"  
17445. .21830

misc\_feature /note="assembly\_fragment"  
21931. .25081

misc\_feature /note="assembly\_fragment"  
25182. .29136  
misc\_feature /note="assembly\_fragment"  
29837. .36653

misc\_feature /note="assembly\_fragment"  
36954. .41270

misc\_feature /note="assembly\_fragment"  
41371. .46765  
misc\_feature /note="assembly\_fragment"  
46866. .57601

misc\_feature /note="assembly\_fragment"  
57702. .65879

misc\_feature /note="assembly\_fragment"  
67980. .76235

misc\_feature /note="assembly\_fragment"  
78336. .89176

misc\_feature /note="assembly\_fragment"  
89277. .106081

misc\_feature /note="assembly\_fragment"  
106182. .121387

misc\_feature /note="assembly\_fragment"  
121488. .137207

misc\_feature /note="assembly\_fragment"  
137308. .153745

misc\_feature /note="assembly\_fragment"  
153846. .172307

BASE COUNT 44249 a 40378 c 41592 g 43880 t  
ORIGIN 2208 others

Query Match 7.7% Score 35.8 DB 71: Length 172307;  
Best Local\_Similarity 52.3% Pred. No. 7.6;  
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Query 214 ggccctgatcgatttcatttccatgtcgatttaccccgatgtggcttccatcaaca 273  
Db 164408 GCTCTGCCCCGGTAGACTTAAAGGATGATGGCCCTGCTGGGACACT 164467  
Query 274 gagatgcccgtatcacccatcccgtgaccaacccgttggcttaataaaa 333

Db 164468 GTGGAACTCTGGATCCCTCTAAGGGCTGTGCTACCCCTGATTTGAAACATA 164527  
Qy 334 cctatcttgcatttcggataacttcc 364  
Db 164528 CAGAGGCTCTCCTGCAAGAAACATGCC 164558

RESULT 13  
AL589962/c  
LOCUS 182529 bp DNA  
DEFINITION Homo sapiens chromosome 6 clone RP11-233K4, \*\*\* SEQUENCING IN  
PROGRESS \*\*, 6 unordered pieces.  
ACCESSION AL589962  
VERSION AL589962.1 GI:13336817  
KEYWORDS HTGS\_PHASE1; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
ORGANISM Homo sapiens  
SOURCE human  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 182529)  
AUTHORS Almeida,J.  
TITLE Direct Submission  
JOURNAL Submitted (15-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail: humquery@sanger.ac.uk Clone  
requests: clonerequest@sanger.ac.uk  
COMMENT ----- Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: bA233K4  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Sequencing vector: M13; M7815; 28% of reads  
Sequencing vector: plasmid; M08752; 71% of reads  
Chemistry: Dye-terminator Big Dye; 73% of reads  
Chemistry: Dye-primer Big Dye; 0% of reads  
Consensus quality: 181426 bases at least Q40  
Consensus quality: 181607 bases at least Q30  
Consensus quality: 181768 bases at least Q20  
Insert size: 182029; sum-of-contigs  
Insert size: 168064; 10.5% error; agarose-fp  
Quality coverage: 14.10x in Q20 bases; agarose-fp  
coverage: 15.93x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 81014: contig of 81014 bp in length  
\* 81015 81114: gap of 100 bp  
\* 81115 105285: contig of 24711 bp in length  
\* 105286 105385: gap of 100 bp  
\* 105386 119420: contig of 14035 bp in length  
\* 119421 119520: gap of 100 bp  
\* 119521 141236: contig of 21716 bp in length  
\* 141237 141336: gap of 100 bp  
\* 141337 180096: contig of 38760 bp in length  
\* 180097 180196: gap of 100 bp  
\* 180197 182529: contig of 2333 bp in length.  
FEATURES  
source  
Location/Qualifiers  
1. .182529  
/organism="Homo sapiens"  
/db\_xref="taxon: 9606"  
/chromosome="6"  
/clone="RP11-233K4"





Search completed: November 7, 2001, 04:14:09  
Job time: 2979 sec

---

Result No.	Score	Query Match	Length	DB ID	Description
1	34	7.3	549	21 AAC95226	Cat flea head and Human head/neck tu
2	32.4	7.0	607	21 AAC69893	Genomic fragment #
c	31.8	6.8	89047	22 AAF2854	Human retrovirus L
c	31.6	6.8	963	21 AAA98219	Polymorphic repeat
5	30.6	6.6	50000	21 AAA96367	M. tuberculosis an
6	30.2	6.5	1166	20 AAZ19410	M. tuberculosis re
c	30.2	6.5	1166	20 AAZ19228	Arabidopsis thalia
c	30	6.5	684	21 AAC52100	Arabidopsis thalia
c	30	6.5	757	21 AAC02266	Arabidopsis thalia
c	30	6.5	1233	21 AAC38681	Arabidopsis thalia
c	30	6.5	1329	21 AAC34844	

The invention relates to novel cat flea (Ctenocephalides felis) nucleic acids which are expressed in hindgut and Malpighian tubule (HMT) tissue or head and nerve cord (HNC) tissue. The invention also relates to the encoded proteins. The invention additionally encompasses expression constructs, recombinant viruses and recombinant cells comprising the nucleic acids of the invention, recombinant production of the proteins, antibodies against the proteins, a method of identifying inhibitors of the proteins, and compositions comprising the proteins they encode may be used in the prevention, treatment and diagnosis of diseases associated with flea infestations. For example, the nucleic acids may be used to produce an HMT or HNC protein according to standard recombinant DNA methodology by inserting the nucleic acids into a host cell and culturing the cell to express the protein. The HMT and HNC nucleic acids may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect and quantitate the presence of cat flea or other homologous nucleic acid sequences in samples. They may also be used to study the expression and function of the proteins and their role in metabolism. The HMT and HNC proteins may be used as antigens in the production of specific antibodies, and in assays to identify modulators (agonists and antagonists) of HMT and/or HNC protein expression and activity. The anti-HMT/HNC protein antibodies and antagonists may also be used to downregulate protein expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of flea polypeptides in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The present sequence represents a cat flea HNC cDNA of the invention.

Query Match	Best Local Similarity	Score	DB 21:	Length
93	7.3%	34		549
	Best Local Matches	Pred. No.	0.13;	
	37	0		
	Conservative	Mismatches	5	
		Indels	0	
93	9ccgcgttacacggcgactgagaaaacctgtgttacccgtctaa	134		
508	9ccgcgttacacggcgactgtggaaacctgtgttacccgtctaa	549		



CC. method for determining the predisposition of a human subject to develop CC. meiotic autoimmune disease. The method comprises detecting a PMR sequence in the CC. CD28, ICOS gene or CTLA4 gene of the human costimulatory receptor gene CC. locus (hCRL). PMR sequences vary in length among individuals and can be CC. amplified to generate products that differ in size. These products can CC. then be detected by rapid and convenient high resolution processes. The CC. method is useful for determining the predisposition of insulin-dependent CC. diabetes mellitus (IDDM), Addison's disease, Graves disease, autoimmune CC. hypothyroidism, myasthenia gravis, thymoma, lupus, thyroiditis, CC. postpartum thyroiditis, rheumatoid arthritis, Hashimoto's disease, CC. coeliac disease and leprosy. PMR sequences within hCRL are useful as CC. markers in a variety of assays and in the field of forensic medicine, CC. disease diagnosis and human genome mapping.



PR	19-JUL-1999;	990US-0144332.	PR	13-OCT-1999;	990US-0159294.
PR	19-JUL-1999;	990US-0144333.	PR	13-OCT-1999;	990US-0159295.
PR	19-JUL-1999;	990US-0144334.	PR	14-OCT-1999;	990US-0159299.
PR	19-JUL-1999;	990US-0144335.	PR	14-OCT-1999;	990US-0159330.
PR	20-JUL-1999;	990US-0144357.	PR	14-OCT-1999;	990US-0159331.
PR	20-JUL-1999;	990US-0144622.	PR	14-OCT-1999;	990US-0159637.
PR	21-JUL-1999;	990US-0144884.	PR	14-OCT-1999;	990US-0159638.
PR	21-JUL-1999;	990US-0145086.	PR	18-OCT-1999;	990US-0159584.
PR	21-JUL-1999;	990US-0145088.	PR	21-OCT-1999;	990US-0160741.
PR	22-JUL-1999;	990US-0145085.	PR	21-OCT-1999;	990US-0160767.
PR	22-JUL-1999;	990US-0145087.	PR	21-OCT-1999;	990US-0160778.
PR	22-JUL-1999;	990US-0145089.	PR	21-OCT-1999;	990US-0160770.
PR	22-JUL-1999;	990US-0145194.	PR	21-OCT-1999;	990US-0160814.
PR	23-JUL-1999;	990US-0145145.	PR	21-OCT-1999;	990US-0160815.
PR	23-JUL-1999;	990US-0145218.	PR	22-OCT-1999;	990US-0160980.
PR	23-JUL-1999;	990US-0145224.	PR	22-OCT-1999;	990US-0160981.
PR	26-JUL-1999;	990US-0145276.	PR	22-OCT-1999;	990US-0160989.
PR	27-JUL-1999;	990US-0145913.	PR	25-OCT-1999;	990US-0161404.
PR	27-JUL-1999;	990US-0145918.	PR	25-OCT-1999;	990US-0161405.
PR	27-JUL-1999;	990US-0145919.	PR	26-OCT-1999;	990US-0161559.
PR	28-JUL-1999;	990US-0145951.	PR	26-OCT-1999;	990US-0161360.
P/	02-AUG-1999;	990US-0146386.	PR	26-OCT-1999;	990US-0161361.
PR	02-AUG-1999;	990US-0146388.	PR	28-OCT-1999;	990US-0161920.
PR	02-AUG-1999;	990US-0146389.	PR	28-OCT-1999;	990US-0161992.
PR	03-AUG-1999;	990US-0147038.	PR	28-OCT-1999;	990US-0161993.
PR	04-AUG-1999;	990US-0147204.	PR	29-OCT-1999;	990US-0162142.
PR	04-AUG-1999;	990US-0147300.	Query	6	Score 30;
PR	05-AUG-1999;	990US-0147192.	Match	53 4;	Pred. No. 3 3;
PR	05-AUG-1999;	990US-0147260.	Best Local	Matches 63;	Mismatches 55;
PR	06-AUG-1999;	990US-0147303.	Similarity	0;	Indels 0;
PR	06-AUG-1999;	990US-0147416.	Matches	63;	Gaps 0;
PR	09-AUG-1999;	990US-0147433.	Qy	165	gtatctcaatgaccacatgcctccaaaccatgcacctaattggccatgtatcg 224
PR	09-AUG-1999;	990US-0147935.	Db	169	GTTTCCGATGCCAACCTCCGACTCTCGATCCAAACCATACCCCTCATAGATCT 110
PR	10-AUG-1999;	990US-0148171.	Qy	225	gatttcttcattatgtcgatttaccatgtccatataacataggatgtcc 282
PR	11-AUG-1999;	990US-0148319.	Db	109	AAAATCTCCGAAATTCCGAAATCCTCCGTCACGGAGATCCACGCCGTTGAAAGCC 52
PR	12-AUG-1999;	990US-0148311.	RESULT	9	
PR	13-AUG-1999;	990US-0148565.	ID	AAA02266	standard; cDNA; 757 BP.
PR	13-AUG-1999;	990US-0148634.	XX	XX	
PR	16-AUG-1999;	990US-0149388.	AC	AAA02266;	
PR	17-AUG-1999;	990US-0149175.	XX	XX	
PR	18-AUG-1999;	990US-0149467.	DT	19-MAY-2000	(first entry)
PR	20-AUG-1999;	990US-0149722.	XX	XX	
PR	20-AUG-1999;	990US-0149733.	DE	Human	colon cancer cell line polynucleotide sequence SEQ ID NO:2257.
PR	23-AUG-1999;	990US-0149929.	XX	KW	Human; colon cancer; tumour; diagnosis; gene expression product;
PR	23-AUG-1999;	990US-0149930.	XX	KW	probe; detection; cancerous state; metastasis; identification;
PR	25-AUG-1999;	990US-0150566.	XX	KW	breast cancer; oestrogen receptor-positive breast cancer; therapy;
PR	26-AUG-1999;	990US-0150884.	XX	KW	oestrogen receptor-negative breast cancer; lung cancer; ss.
PR	27-AUG-1999;	990US-0151063.	XX	XX	
F	10-SEP-1999;	990US-0151066.	OS	OS	Homo sapiens
PR	13-SEP-1999;	990US-0151378.	XX	XX	
PR	15-SEP-1999;	990US-0151309.	PN	PN	W09958675-A2.
PR	16-SEP-1999;	990US-0154039.	XX	XX	
PR	20-SEP-1999;	990US-0154779.	PD	18-NOV-1999.	
PR	22-SEP-1999;	990US-0155139.	XX	XX	
PR	23-SEP-1999;	990US-0155236.	PF	13-MAY-1999;	99WO-US10602.
PR	07-SEP-1999;	990US-0155106.	XX	XX	
PR	10-SEP-1999;	990US-0155307.	PR	14-MAY-1998;	98US-0085426.
PR	13-SEP-1999;	990US-0155388.	XX	XX	
PR	15-SEP-1999;	990US-0156458.	PR	15-MAY-1998;	98US-0085537.
PR	16-SEP-1999;	990US-0156596.	XX	XX	
PR	04-OCT-1999;	990US-0157117.	PR	15-MAY-1998;	98US-0085936.
PR	05-OCT-1999;	990US-0157753.	PR	21-OCT-1998;	98US-0105244.
PR	06-OCT-1999;	990US-0157865.	PR	27-OCT-1998;	98US-0105877.
PR	07-OCT-1999;	990US-0158029.	XX	XX	
PR	08-OCT-1999;	990US-0158232.	PA	(CHIR ) CHIRON CORP.	
PR	12-OCT-1999;	990US-0158369.	PA	(HYSE- ) HYSEQ INC.	
PR	13-OCT-1999;	990US-0159233.	XX	XX	

PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J; PR 16-APR-1999; 99US-0129845; 99US-0130077; 99US-0130077; 99US-0130449; 99US-0130510; 99US-0130891; 99US-0131449; 99US-0132048; 99US-0132407; 99US-0132484; 99US-0132485; 99US-0132486; 99US-0132487; 99US-0132863; 99US-0134256; 99US-0134218; 99US-0134219; 99US-0134221; 99US-0134370; 99US-0134768; 99US-0134941; 99US-0135124; 99US-0135353; 99US-0135629; 99US-0136021; 99US-0136392; 99US-0136782; 99US-0137222; 99US-0137528; 99US-0137724; 99US-0137702; 99US-0138094; 99US-0138540; 99US-0138847; 99US-0139119; 99US-0139452; 99US-0139453; 99US-0139452; 99US-0139454; 99US-0139454; 99US-0139455; 99US-0139456; 99US-0139456; 99US-0139457; 99US-0139458; 99US-0139458; 99US-0139460; 99US-0139461; 99US-0139462; 99US-0139463; 99US-0139750; 99US-0139763; 99US-0139817; 99US-0139899; 99US-0139899; 99US-0140253; 99US-0140354; 99US-0140695; 99US-0142390; 99US-0142803; 99US-0142920; 99US-0142977; 99US-0143542; 99US-0143624; 99US-0144005; 99US-0144085; 99US-0144086; 99US-0144325; 99US-0144331; 99US-0144332; 99US-0144333.

XX Claim 1: Page 886; 1097pp; English.

XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA libraries constructed from human colon cancer cell lines. The present invention also describes a method of detecting differentially expressed genes correlated with a cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample derived from a cell suspected of being cancerous, where detection of the differentially expressed gene product is correlated with a cancerous state of the cell from which the test sample was derived.

CC The polynucleotides sequences can be used in a method for detecting differentially expressed genes correlated with a cancerous state of a mammalian cell. The polynucleotides can also be used as probes for detecting and mapping related genes. They can be used in diagnosis and prognosis of diseases and disorders (e.g. identification of pre-metastatic or metastatic cancerous states, stages of cancer, or responsiveness of cancer to therapy). This is particularly for breast cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-negative breast cancer, lung cancer, and colon cancer.

XX Sequence 757 BP; 155 A; 205 C; 168 G; 199 T; 30 other;

XX SO Query Match 6.58; Score 30; DB 21; Length 757; Best Local Similarity 68.4%; Pred. No. 3.4; Matches 39; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 223 cggattttcttaccatgtcgatattaccctatggcgcttcataacaaggaggat 279  
Db 700 cggattttcttaccatgtcgatattaccatggcgacttcataatgtt 756

RESULT 10 AAC38681/C ID AAC38681 standard; DNA; 1233 BP.

X Ac AAC38681;  
XX DR 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 21860.  
XX KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX OS Arabidopsis thaliana.  
XX EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PN 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-01223180.  
PR 09-MAR-1999; 99US-0123518.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126735.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128224.  
PR 08-APR-1999; 99US-0128714.



PR	23-APR-1999;	990US-0130510.
PR	23-APR-1999;	990US-0130891.
PR	28-APR-1999;	990US-0131449.
PR	30-APR-1999;	990US-0131048.
PR	30-APR-1999;	990US-013407.
PR	04-MAY-1999;	990US-0132484.
PR	05-MAY-1999;	990US-0132485.
PR	06-MAY-1999;	990US-0132486.
PR	06-MAY-1999;	990US-0132487.
PR	07-MAY-1999;	990US-0132863.
PR	11-MAY-1999;	990US-0134256.
PR	14-MAY-1999;	990US-0134218.
PR	14-MAY-1999;	990US-0134219.
PR	14-MAY-1999;	990US-0134221.
PR	14-MAY-1999;	990US-0134270.
PR	18-MAY-1999;	990US-0134768.
PR	19-MAY-1999;	990US-0134769.
PR	01-JUN-1999;	990US-013124.
PR	21-MAY-1999;	990US-0135353.
PR	24-MAY-1999;	990US-0135629.
PR	25-MAY-1999;	990US-0136021.
PR	27-MAY-1999;	990US-0136392.
PR	28-MAY-1999;	990US-0136782.
PR	01-JUN-1999;	990US-013722.
PR	03-JUN-1999;	990US-0137528.
PR	04-JUN-1999;	990US-0131502.
PR	07-JUN-1999;	990US-0137724.
PR	08-JUN-1999;	990US-0136094.
PR	10-JUN-1999;	990US-0136154.
PR	10-JUN-1999;	990US-0138847.
PR	14-JUN-1999;	990US-0139119.
PR	16-JUN-1999;	990US-0139452.
PR	16-JUN-1999;	990US-0139453.
PR	17-JUN-1999;	990US-0139492.
PR	18-JUN-1999;	990US-0139454.
PR	18-JUN-1999;	990US-0139455.
PR	18-JUN-1999;	990US-0139456.
PR	18-JUN-1999;	990US-0139457.
PR	18-JUN-1999;	990US-0139458.
PR	18-JUN-1999;	990US-0139459.
PR	21-JUN-1999;	990US-0139460.
PR	18-JUN-1999;	990US-0139461.
PR	18-JUN-1999;	990US-0139462.
PR	23-JUN-1999;	990US-0139463.
PR	24-JUN-1999;	990US-0139465.
PR	28-JUN-1999;	990US-0139763.
PR	29-JUN-1999;	990US-0139817.
PR	30-JUN-1999;	990US-0139899.
PR	01-JUL-1999;	990US-0130353.
PR	01-JUL-1999;	990US-0120354.
PR	02-JUL-1999;	990US-0120695.
PR	06-JUL-1999;	990US-0120823.
PR	09-JUL-1999;	990US-0120991.
PR	12-JUL-1999;	990US-0121287.
PR	12-JUL-1999;	990US-0121842.
PR	13-JUL-1999;	990US-012154.
PR	14-JUL-1999;	990US-012055.
PR	15-JUL-1999;	990US-012390.
PR	16-JUL-1999;	990US-012803.
PR	16-JUL-1999;	990US-012920.
PR	16-JUL-1999;	990US-014085.
PR	16-JUL-1999;	990US-0144086.
PR	19-JUL-1999;	990US-014325.
PR	19-JUL-1999;	990US-014331.
PR	19-JUL-1999;	990US-014332.
PR	19-JUL-1999;	990US-014333.
PR	19-JUL-1999;	990US-014334.
PR	20-JUL-1999;	990US-014352.



PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134688.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145931.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146316.
PR	21-MAY-1999;	99US-0135333.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-0156291.	PR	02-AUG-1999;	99US-0146399.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136921.	PR	04-AUG-1999;	99US-014704.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-014702.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137228.	PR	05-AUG-1999;	99US-0147199.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147260.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147303.
PR	08-JUN-1999;	99US-0138094.	PR	06-AUG-1999;	99US-0147416.
PR	09-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138847.	PR	09-AUG-1999;	99US-0147495.
PR	14-JUN-1999;	99US-0139119.	PR	10-AUG-1999;	99US-0148171.
PR	16-JUN-1999;	99US-0139528.	PR	11-AUG-1999;	99US-0148319.
PR	18-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148365.
PR	18-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149939.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	22-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	23-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150984.
PR	23-JUN-1999;	99US-0139464.	PR	27-AUG-1999;	99US-0151065.
PR	24-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151066.
PR	28-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151080.
PR	21-JUN-1999;	99US-0139817.	PR	30-AUG-1999;	99US-0151103.
PR	30-JUN-1999;	99US-0139899.	PR	31-AUG-1999;	99US-0151138.
PR	01-JUL-1999;	99US-0140153.	PR	01-SEP-1999;	99US-0151930.
PR	12-JUL-1999;	99US-0140354.	PR	07-SEP-1999;	99US-015263.
PR	13-JUL-1999;	99US-0140354.	PR	10-SEP-1999;	99US-0152970.
PR	01-JUL-1999;	99US-01215.	PR	13-SEP-1999;	99US-0153758.
PR	28-JUN-1999;	99US-0140823.	PR	15-SEP-1999;	99US-0154018.
PR	29-JUN-1999;	99US-0140991.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0141842.	PR	20-SEP-1999;	99US-0154779.
PR	01-JUL-1999;	99US-014215.	PR	22-SEP-1999;	99US-0155139.
PR	02-JUL-1999;	99US-0142055.	PR	23-SEP-1999;	99US-015586.
PR	06-JUL-1999;	99US-0142390.	PR	24-SEP-1999;	99US-0155559.
PR	08-JUL-1999;	99US-0142803.	PR	28-SEP-1999;	99US-0156458.
PR	09-JUL-1999;	99US-0142220.	PR	29-SEP-1999;	99US-0155956.
PR	12-JUL-1999;	99US-0142977.	PR	04-OCT-1999;	99US-0156317.
PR	13-JUL-1999;	99US-014354.	PR	13-OCT-1999;	99US-0156367.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0156293.
PR	15-JUL-1999;	99US-0144005.	PR	13-OCT-1999;	99US-0156294.
PR	19-JUL-1999;	99US-014405.	PR	13-OCT-1999;	99US-0156295.
PR	16-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0156329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0156330.
PR	20-JUL-1999;	99US-0144335.	PR	21-OCT-1999;	99US-0160767.
PR	19-JUL-1999;	99US-0144332.	PR	21-OCT-1999;	99US-0160768.
PR	20-JUL-1999;	99US-0144632.	PR	21-OCT-1999;	99US-0160770.
PR	19-JUL-1999;	99US-014484.	PR	21-OCT-1999;	99US-0160814.
PR	21-JUL-1999;	99US-0144814.	PR	21-OCT-1999;	99US-0160815.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160980.
PR	22-JUL-1999;	99US-0145085.	PR	22-OCT-1999;	99US-0160981.
PR	22-JUL-1999;	99US-0145087.	PR	22-OCT-1999;	99US-0160982.
PR	22-JUL-1999;	99US-0145192.	PR	22-OCT-1999;	99US-0160983.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160984.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160985.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.



XX  
PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;  
XX  
DR WPI; 2000-656323/63.

XX  
PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic  
PT acids useful for the prevention, diagnosis and treatment of flea  
PT infestations -

XX  
PS Claim 26; Page 784; 964pp; English.

XX  
The invention relates to novel cat flea (*Ctenocephalides felis*) nucleic  
CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue  
CC or head and nerve cord (HNC) tissue. The invention also relates to the  
CC encoded proteins. The invention also encompasses expression  
CC constructs, recombinant viruses and recombinant cells comprising the  
CC nucleic acids of the invention, recombinant production of the proteins,  
CC antibodies against the proteins, a method of identifying inhibitors of  
CL the proteins, and compositions comprising the inhibitors for  
CC administration to an animal. The nucleic acids, and the proteins they  
CC encode may be used in the prevention, treatment and diagnosis of diseases  
CC associated with flea infestations. For example, the nucleic acids may be  
CC used to produce an HMT or HNC protein according to standard recombinant  
CC DNA methodology by inserting the nucleic acids into a host cell and  
CC culturing the cell to express the protein. The HMT and HNC nucleic acids  
CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect  
CC and quantitate the presence of cat flea or other homologous nucleic acid  
CC sequences in samples. They may also be used to study the expression and  
CC function of the proteins and their role in metabolism. The HMT and HNC  
CC proteins may be used as antigens in the production of specific  
CC antibodies, and in assays to identify modulators (agonists and  
CC antagonists) of HMT and/or HNC protein expression and activity. The  
CC anti-HMT/HNC protein antibodies and antagonists may also be used to  
CC downregulate protein expression and activity. The antibodies may also be  
CC used as diagnostic agents for detecting the presence of flea polypeptides  
CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA). The  
CC present sequence represents a cat flea HMT cDNA of the invention.

XX  
SQ Sequence 640 BP; 208 A; 135 C; 133 G; 153 T; 11 other;

Query Match 6.3%; Score 29.4; DB 21; Length 640;  
Best Local Similarity 53.3%; Pred. No. 5.1;  
Matches 96; Conservative 0; Mismatches 80; Indels 4; Gaps 2;  
Q: 37 agactcgctaaagacccgtccaaatccctttatggcgtaatgcactggcg 96  
Db 452 aactcgaggggggccgtccaaatcgcttatgtgatcgatataactcgcg 511  
Qy 97 cgcttacacg -cggtctgadaaaacctgtcgtaaccctctaattttcacccatccctat 154  
Db 512 cgtttacacgctgtgtggaaaccctgtcgtaaccctatgtgnacatccctt 571  
Qy 155 cgactggcggtatctcaatgaccaccatcgcttcaaccatggcaactatggcaatg 214  
Db 572 cncatcg -cgatataaaaggcccccncatgncatgtggaaatgtggaaang 629

Search completed: November 7, 2001, 04:11:03  
Job time: 2723 sec

```

44: em_esthun10:*
45: em_esthun11:*
46: em_esthun12:*
47: em_esthun13:*
48: em_esthun14:*
49: em_esthun15:*
50: em_esthun16:*
51: em_esthun17:*
52: em_esthun18:*
53: em_esthun19:*
54: em_esthun20:*
55: em_esthun21:*
56: em_esthun22:*
57: em_esthun23:*
58: em_esthun24:*
59: em_esthun25:*
60: em_esthun26:*
61: em_esthun27:*
62: em_esthun28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estp11:*
73: em_estp12:*
74: em_estp13:*
75: em_estp14:*
76: em_estp15:*
77: em_estp16:*
78: em_estp17:*
79: em_estp18:*
80: em_estp19:*
81: em_estp110:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_cstro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: em_est30:*
103: em_est31:*
104: em_est32:*
105: em_est33:*
106: em_est34:*
107: em_est35:*
108: em_est36:*
109: em_est37:*
110: em_est38:*
111: em_est39:*
112: em_est40:*
113: em_est41:*
114: em_est42:*
115: em_est43:*
116: em_est44:*
117: em_est45:*
118: em_est46:*
119: em_est47:*

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2001, 03:23:35 ; Search time 1170.27 Seconds  
(without alignments)  
3756.037 Million cell updates/sec

Title: US-09-521-640-2  
Sequence: 1 ggggatcgatcgatccac.....cccttctgttcaaaacacn 465

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Scored: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: qb\_est1:\*
2: qb\_est2:\*
3: qb\_est3:\*
4: qb\_est4:\*
5: qb\_est5:\*
6: qb\_est6:\*
7: qb\_est7:\*
8: qb\_est8:\*
9: qb\_est9:\*
10: qb\_est10:\*
11: qb\_est11:\*
12: qb\_est12:\*
13: qb\_est13:\*
14: qb\_est14:\*
15: qb\_est15:\*
16: qb\_est16:\*
17: qb\_est17:\*
18: qb\_est18:\*
19: qb\_est19:\*
20: qb\_est20:\*
21: qb\_est21:\*
22: qb\_est22:\*
23: qb\_est23:\*
24: qb\_est24:\*
25: qb\_est33:\*
26: qb\_est34:\*
27: qb\_est35:\*
28: qb\_est36:\*
29: qb\_est37:\*
30: qb\_est38:\*
31: qb\_est39:\*
32: qb\_est40:\*
33: em\_estba:\*
34: em\_estfun:\*
35: em\_esthun1:\*
36: em\_esthun2:\*
37: em\_esthun3:\*
38: em\_esthun4:\*
39: em\_esthun5:\*
40: em\_esthun6:\*
41: em\_esthun7:\*
42: em\_esthun8:\*
43: em\_esthun9:\*

```

117: qb_est48;*
118: qb_est49;*
119: qb_est50;*
120: qb_est51;*
121: qb_est52;*
122: qb_est53;*
123: qb_est54;*
124: qb_est55;*
125: qb_est56;*
126: qb_est57;*
127: qb_est58;*
128: qb_est59;*
129: qb_est60;*
130: qb_est61;*
131: qb_est62;*
132: qb_est63;*
133: qb_est64;*
134: qb_est65;*
135: qb_est66;*
136: qb_est67;*
137: qb_est76;*
138: qb_est77;*
139: qb_est78;*
140: qb_est79;*
141: qb_est80;*
142: qb_est81;*
143: qb_est82;*
144: qb_est83;*
145: qb_est84;*
146: qb_est85;*
147: qb_est86;*
148: qb_est87;*
149: qb_est88;*
150: qb_est89;*
151: qb_est90;*
152: qb_est99;*
153: qb_est100;*
154: qb_est101;*
155: qb_est102;*
156: qb_est103;*
157: qb_est104;*
158: qb_est105;*
159: qb_est106;*
160: qb_est67;*
161: qb_est68;*
162: qb_est69;*
163: qb_est70;*
164: qb_est71;*
165: qb_est72;*
166: qb_est73;*
167: qb_est74;*
168: qb_est31;*
169: qb_est92;*
170: qb_est93;*
171: qb_est9;*
172: qb_est95;*
173: qb_est96;*
174: qb_est97;*
175: qb_est98;*
176: em_esthmu29;*
177: em_esthmu30;*
178: em_esthmu31;*
179: em_esthmu32;*
180: em_esthmu33;*
181: em_esthmu3;*
182: em_estp11;*
183: em_estro21;*
184: em_estro22;*
185: em_estro23;*
186: em_htc;*
187: qb_est107;*
188: qb_est108;*
189: qb_est109;*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length	DB ID	Description		
1	89.2	19.2	579	225	AQ210729 HS_2229_A		
2	64	13.8	489	233	AQ761577 HS_3187_A		
3	63.8	13.7	468	233	AQ806004 HS_3025_A		
4	60.6	13.5	740	225	AQ914869 HS_3220_A		
5	53.6	11.5	397	225	AQ208121 HS_3229_B		
6	50.2	10.8	1195	162	BE035400 MO_0410_M		
7	45.6	9.8	677	32	AV702190 AV702190		
8	45.2	9.7	173	226	AQ304554 HS_3247_B		
9	44.4	9.5	300	167	BE401707 CNW02EL02		
10	43.6	9.4	205	233	AQ761515 HS_3161_A		
11	43.2	9.3	736	32	AV721604 AV721604		
12	43	9.2	946	32	AV726956 AV726956		
13	42.8	9.2	168	234	AQ823152 HS_3186_B		
14	42.6	9.2	659	32	AV705287 AV705287		
C	15	42.4	9.1	312	256 B45121		
C	16	42.2	9.1	259	AQ75293 HS_3151_A		
C	17	42.2	9.1	1101	219 CNS0012		
C	18	41.4	8.9	419	224 AQ116709		
C	19	41.4	8.9	818	32 AV726386 AV726386		
C	20	41	8.8	545	139 BE755460		
C	21	40.6	8.7	982	15 A1068596		
22	40.4	8.7	403	233	AQ783487 HS_3063_A		
23	40.4	8.7	714	104	A1965418 sc71607_Y		
24	40.4	8.7	847	32	AV726234 AV726234		
25	40.2	8.6	981	15	A1068440 mge0002b		
26	40	8.6	146	235 AQ891463	AQ891463 HS_3111_B		
27	40	8.6	432	224 AQ139216	AQ139216 HS_3090_A		
28	40	8.6	961	152 BG343702	BG343702 HVSMB9000		
29	39.6	8.5	803	112 AW155132	AW155132 mge00020		
30	39.6	8.5	870	32 AV726543	AV726543 AV726543		
31	39.4	8.5	920	32 AV772728	AV772728 AV772728		
32	39.2	8.4	819	153 BG418812	BG418812 HVSMB002		
33	39	8.4	768	15 A1068832	A1068832 mge0004b		
34	38.8	8.3	457	233 AQ775019	AQ775019 HS_3155_A		
35	38.8	8.3	890	164 BE214026	BE214026 HV_CSB000		
36	38.8	8.3	985	162 BE034546	BE034546 MK01B09_M		
37	38.6	8.3	1200	110 AV761334	AV761334 AV761384		
38	38	8.2	929	152 BG343386	BG343386 HVSMB9000		
39	38	8.2	972	106 AL576318	AL576318 AL576318		
40	37.8	8.1	287	233 AQ805901	AQ805901 HS_3335_A		
41	37.8	8.1	1017	162 BE039590	BE039590 QC02804_O		
42	37.6	8.1	430	226 AQ301842	AQ301842 HS_3195_A		
43	37.6	8.1	945	164 BE195597	BE195597 HVSMB009		
44	36.8	7.9	263	120 AW800405	AW800405 MR22-DN006		
45	36.8	7.9	802	32 AV725250	AV725250 AV725250		

## ALIGNMENTS

RESULT	1	1	579 bp	DNA	GSS	27-JUL-1999	
ORGANISM	Homo sapiens	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	DEFINITION	HS_2229_A1_C03_MR C1T Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2229 Col=5 Row=E, DNA sequence.	COMMENT	Human	
REFERENCE	1 (bases 1 to 579)	Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and Hood, L.		1 (bases 1 to 489)	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)	JOURNAL	scanning the human genome
AUTHORS	1	Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and Hood, L.		1	993805f.	MEDLINE	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
TITLE	1	Sequence-tagged connectors: A sequence approach to mapping and sequencing the human genome		1	993805g.	COMMENT	High throughput Sequencing Center
RESULT	1	Sequence-tagged connectors: A sequence approach to mapping and sequencing the human genome		1	993805h.	ORGANISM	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
LOCUS	AQ210729	Human		1	993805i.	COMMENT	High throughput Sequencing Center
DEFINITION	HS_2229_A1_C03_MR C1T Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2229 Col=5 Row=E, DNA sequence.	Human		1	993805j.	ORGANISM	Human
ACCESSION	AQ210729	Human		1	993805k.	COMMENT	High throughput Sequencing Center
VERSION	1	Human		1	993805l.	ORGANISM	Human
KEYWORDS	GSS.	Human		1	993805m.	COMMENT	High throughput Sequencing Center
SOURCE	GSS.	Human		1	993805n.	ORGANISM	Human
ORGANISM	Homo sapiens	Human		1	993805o.	COMMENT	High throughput Sequencing Center
DEFINITION	Human	Human		1	993805p.	ORGANISM	Human
COMMENT	Human	Human		1	993805q.	COMMENT	High throughput Sequencing Center
CONTACT	Wallace, J.C., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and Hood, L.	Human		1	993805r.	ORGANISM	Human
UNIVERSITY	University of Washington	Human		1	993805s.	COMMENT	High throughput Sequencing Center
401 QUEEN AVE NORTH, SEATTLE, WA 98109, USA	401 Queen Ave North, Seattle, WA 98109, USA	Human		1	993805t.	ORGANISM	Human
TELEPHONE: (206) 516-3618	TELEPHONE: (206) 516-3618	Human		1	993805u.	COMMENT	High throughput Sequencing Center



source	<p>1. :740  <i>/organism="Oryza sativa"  /strain="Japonica"  /cultivar="Nipponbare"  /db_xref="taxon:4330"  /clone="rbab005020R"  /clone.lib="CUGI Rice BAC Library (ECORI)"  /tissue_type="Leaf"  /lab_host="E. coli DH10B"  /NOTE: Vector: pBACIndigo; Site_1: EcoRI; Site_2: EcoRI;  Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate.  Monotonydonus rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed BAC library from <i>Oryza sativa</i>, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation of a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (<a href="http://www.genome.clemson.edu/">www.genome.clemson.edu</a>)."</i></p>
BASE COUNT	195 a
ORIGIN	173 c 190 g 181 t 1 others
FEATURES	<p>source  1. .397  1. Location/Qualifiers  1. .397  1. /organism="Homo sapiens"  1. /db_xref="taxon:9606"  1. /clone.lib="CIT Approved Human Genomic Sperm Library D"  1. /sex="male"  1. /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E. coli DH10B"</p>
BASE COUNT	93 a
ORIGIN	111 c 85 g 106 t 2 others
FEATURES	<p>source  1. .397  1. Location/Qualifiers  1. .397  1. /organism="Homo sapiens"  1. /db_xref="taxon:9606"  1. /clone.lib="CIT Approved Human Genomic Sperm Library D"  1. /sex="male"  1. /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E. coli DH10B"</p>
Query Match	13.0%
Best Local Similarity	61.7%
Matches	113; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
Query	ttccacatcccccatacgactgggtatctcaatggaccatccatcgccctccaaacctg 198
Result	6
Db	342 TCGAGACATCCCCCTTCGAGCTGGTATAGCAAAAGGGACGGATGCCCTACAGTGT 401
DEFINITION	BE035400 1195 bp mRNA Mod10 Mo Mesembryanthemum crystallinum CDNA 5' similar to ribosomal protein 117, mRNA sequence.
ACCESSION	BE035400
VERSION	1 (bases 1 to 1195) EST.
KEYWORDS	Mesembryanthemum crystallinum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
ORGANISM	Common ice plant Mesembryanthemum crystallinum Bohnert, H.J., Borch, C., Brazille, S., Brooks, J., Eaton, M., Ferrea, H., Kawasaki, S., McCollough, A., Michalowski, C.B., Palacio, C., Scara, G., Wheeler, M., and Zepeda, G.R.
REFERENCE	Unpublished (2000)
AUTHORS	Bohnert, H.J., Borch, C., Brazille, S., Brooks, J., Eaton, M., Ferrea, H., Kawasaki, S., McCollough, A., Michalowski, C.B., Palacio, C., Scara, G., Wheeler, M., and Zepeda, G.R.
TITLE	
JOURNAL	
COMMENT	University of Arizona Bio Sciences West room 513, Tucson, AZ 85721, USA
FEATURES	<p>source  1. .397  1. Location/Qualifiers  1. .397  1. /organism="Mesembryanthemum crystallinum"  1. /db_xref="taxon:3544"  1. /note="MO"  1. /issue_type="apical meristem and leaf primordia"  1. /av_sage="5 weeks"  1. /note="no stress"</p>
RESULT	5
LOCUS	AQ208121 397 bp DNA
DEFINITION	HS_3220_A2_H07_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3220 Col=14 Row=0, DNA sequence.
ACCESSION	AQ208121
VERSION	GI:3620856
KEYWORDS	
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1. (bases 1 to 397) Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and Hood, L.
TITLE	Sequencing-tagged connectors: A sequence approach to mapping and scanning the human genome

BASE COUNT	ORIGIN	305 a	266 c	301 g	304 t	19 others	RESULT	8
Query Match	10.8%	Score 50.2;	DB 162;	Length 1195;				
Best Local Similarity	66.4%	Pred. No. 8	3e-05;					
Matches	0;	Mismatches	43;	Indels	1;	Gaps		
Qy	88	tcaactgcgcgttacacg-cggactgagaaaaacctcgtaatctttcacaca	146					
Db	874	TAATCTGCCTTACGTGTCAGTGGAAACCTGGTACCAATCTTGAGACA	933					
Qy	147	tcccttatgcactggggatctcatgtacccatcgccatccatggccaaat	206					
Db	934	TCCCCTTTGCGTGGCAATCGAACGCTTACAGTGTCAACTGAT	993					
Qy	207	tggcaatggcc	217					
Db	994	TGCTATGGACC	1004					
RESULT	7							
Av-190	AV702190	677 bp	mRNA	EST	08-OCT-2000			
Db..-NITION	AV702190	ADB	Homo sapiens	CDNA clone	ADBCOH01	5', mRNA sequence.		
ACCESSION	AV702190							
	AV702190.1							
KEYWORDS	EST.							
SOURCE	Homo sapiens							
ORGANISM	Mammalia; Eutheria; Primates; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE	Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z., and Han, Z.							
AUTHORS								
TITLE	Homo sapiens cDNA ADB clones							
JOURNAL	Unpublished (2000)							
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P.R. China tel: 86-21-50801919 ext.45 Fax: 86-21-50801922 Email: hanzz@chgc.sh.cn							
FEATURES	Location/Qualifiers							
source	1.. 677 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="ADBCOH01" /clone_lib="ADB" /tissue_type="Adrenal gland" /dev_stage="Adult" /lab_host="SOL" /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XbaI",							
BASE COUNT	166 a	161 c	156 g	194 t				
ORIGIN								
RESULT	9							
BE401707								
LOCUS	BE401707	300 bp	mRNA	EST	21-JUL-2000			
DEFINITION	CNN02E021 TRBC CNN Wheat Powdery Mildew Resistant Library Triticum aestivum CDNA clone CNN02E021, mRNA sequence.							
ACCESSION	BE401707							
VERSION	BE401707.1							
KEYWORDS	EST.							
SOURCE								
ORGANISM	Triticum aestivum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.							
REFERENCE	1 (bases 1 to 300)							

AUTHORS	Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P., Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P., Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y., Peccifoni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M., Sorrells,M., Warburton,M., and Wenzel,G.
TITLE	International Triticeae EST Cooperative (ITEC): Production of Expressed Sequence Tags for Species of the Triticeae
JOURNAL	Unpublished (2000)
COMMENT	Contact: Jia J Key Lab. of Crop Germplasm & Biotechnology Inst. of Crop Germplasm Resources Beijing 100081 PR CHINA Tel: 86 1 62186623 Fax: 86 1 62186629 Email: jizeng@imim.cnc.ac.cn
FEb.	RES
source	Location/Qualifiers 1. 300 /organism="Triticum aestivum" /cultivar="Powdery Mildew Resistant line" /db_xref="Taxon:4565" /clone="CNW02E021" /clone.lib="ITEC CNW Wheat Powdery Mildew Resistant Library" /tissue_type="leaf" /dev_stage="seedling, challenged with powdery mildew strain" /note="Vector: Lambda TriPLEX; Site_1: SfiI; Primer used 5'-TCGAGATCTGGACGAGC-3'; 500 bp average insert size." 5'-TCGAGATCTGGACGAGC-3'; 500 bp average insert size." 6' others
BASE COUNT	80 a 75 c 67 g 72 t 30 n
ORIGIN	Query Match 9.5%; Score 44.4%; DB 167; Length 300; Best Local Similarity 66.5%; Pred. No. 0.038; Matches 105; Conservative 0; Mismatches 50; Indels 3; Gaps 3
RESULT	10
LOCUS	AQ763115 205 bp DNA
DEFINITION	HS-116_A2-D02_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone plate=3161 Col=4 Row=G, DNA sequence.
ACCESSION	AQ763115
VERSION	1
KEYWORDS	GR:5641231
SOURCE	human. Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	(bases 1-205)
AUTHORS	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L 99380589



LOCUS	AV705287	659 bp	mRNA	clone ADBBE08	5'	EST	09-OCT-2000	Seattle, WA 98195, USA
DEFINITION	AV705287	ADB	Homo sapiens	cdNA				Tel: (206) 616-8744
VERSION	AV705287							Fax: (206) 685-7301
KEYWORDS	EST.							Email: kzackron@u.washington.edu
SOURCE	human.							Sequence Tagged Connector
ORGANISM	Homo sapiens							Plate: CT 782 row: N column: 1
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							Class: BAC ends
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							High quality sequence stop: 312.
REFERENCE	1 (bases 1 to 659)							Location/Qualifiers
AUTHORS	Peng, Y., Song, H., Huang, Q., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Zeng, L., Xu, S., Tu, Y., Jia, J., Fu, G., Ren, S., Lu, G., Hu, R., Chen, J., Chen, Z., and Han, Z.							1. .312
TITLE	Homo sapiens cdNA ADB clones							/organism="Homo sapiens"
JOURNAL	Unpublished (2000)							/db_xref="taxon:9606"
CC-BY-NC	Contact: Zeguang Han							/clone="Plate-CT 782 Col=1 Row=N"
	Chinese National Human Genome Center at Shanghai							/clone.lib="CIT Human Genomic sperm Library C"
	351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai							/sex="M"
	20103, P. R. China							/note="Organism: sperm; Vector: pBelobAC11; BAC Clones in
	Tel: 86-21-50801919 (ex. 45)							E-Cv11 DH10B"
	Fax: 86-21-50801922							
	Email: hanzg@hgc.sh.cn							
FEATURES	source							BASE COUNT
	source							80 a 66 C 95 g 70 t 1 others
	source							ORIGIN
	source							
	source							Query Match
	source							Best Local Similarity
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source							94; Conservative
	source							0; Mismatches
	source							87; Indels 0; Gaps 0;
	source							Matches
	source					</td		



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

**BLACK BORDERS**

**IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

**FADED TEXT OR DRAWING**

**BLURRED OR ILLEGIBLE TEXT OR DRAWING**

**SKEWED/SLANTED IMAGES**

**COLOR OR BLACK AND WHITE PHOTOGRAPHS**

**GRAY SCALE DOCUMENTS**

**LINES OR MARKS ON ORIGINAL DOCUMENT**

**REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

**OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**